Game-Based Human-System Analysis for National Security R&D

Matthew P. Daggett, Timothy J. Dasey, Adam S. Norige, and Robert M. Seater

Serious games are influencing efforts to improve education, health care, defense, and awareness of societal issues by applying gamification to help users develop understanding and skills in these fields and to elicit knowledge from expert users. Researchers at Lincoln Laboratory are transforming traditional research and development processes by using games to design, engineer, and assess more efficient and effective sociotechnical systems for national security needs.

Since the 1950s, MIT Lincoln Laboratory has conducted rigorous systems analysis, full-system prototyping, and development of long-term advanced technologies for national security applications. As the discrete systems of earlier decades have been replaced with complex interconnected systems of systems, traditional modeling and simulation and systems analysis often insufficiently account for human dynamics. These limitations become further exacerbated as a long-standing paradigm of systems as subordinate to operators is being replaced with collaborative workflows enabled by automation and artificial intelligence. To address these challenges, researchers at the Laboratory have developed methodologies and technologies for designing, building, and employing serious games that measure human decision making and that serve as systems analysis tools to assess and facilitate complex human-system dynamics that approximate those of realistic sociotechnical systems. These serious games are a unique tool in the research and development (R&D) process that overcomes the limitations of other methods.

Games are a structured form of play, usually undertaken for enjoyment, achievement, or reward. They have been recorded as part of cultures dating back to the 26th century BCE and are thought to be universal to the human experience. Games have also long been used for U.S. national security purposes, with some of the earliest wargame efforts in the 1800s at the Naval War College. Many early wargames were large tabletop or seminar-style formats used for developing war plans and exercising decision making. With the advent of modern computing, these games have grown into large, sophisticated, distributed semiautomated force simulations, largely focused on informing military training and doctrine.

In the 1970s, a definition of the term serious games emerged to broadly define games as a means to achieve an explicit purpose other than amusement. Under this rubric, gamification has been employed in education, scientific exploration, health care, emergency management, and more. While the use of games in a national security context is often synonymous with wargaming, Lincoln Laboratory's research into games aligns with the broader view of the application of serious games.

Gaming at Lincoln Laboratory

Since 2001, the Laboratory has been developing purpose-built serious games and applying them to its core R&D processes across a variety of mission areas, including air and missile defense; intelligence, surveillance, and reconnaissance; chemical and biological defense; air traffic management; cyber security operations; and emergency response. While built with different objectives, these games fall mainly into three common game applications:

1. Experiential learning. Games are a natural fit for training and can model situations that are rare in practice, dangerous to rehearse, or potentially possible in the future. Inside a virtual environment, a participant can experiment with high-stakes situations in a low-stakes environment, building intuition and mental models for how the environment reacts. Virtual training environments are prominent in training pilots or power plant operators, for whom live training on real equipment is expensive and at risk for catastrophic mistakes. Game-based training may be high-fidelity recreations of the physical world, but they can also be unscripted, abstract, and open-ended experiences, while still focusing on key aspects of complex tasks.

   The Laboratory has applied training games to several domains, including emergency response to improvised nuclear device detonations and management of delays in air traffic systems during severe weather. A common characteristic of all these games was a focus on only an important slice of the problem rather than on a model of the entire problem space. As a result, the games required just minutes to play a scenario, allowing players to engage in many iterations of a scenario in a single sitting. The Laboratory's methodology that combines repetition of short focused experiences with engagement in longer more detailed experiences has proven an effective approach to cover the full spectrum of complex tasks.
2. Concept exploration and requirement analysis. Predicting what technologies will be useful and impactful prior to building a prototype is error prone and can result in expensive redesigns when operators reject the technology at late stages of a development process. Consultation with experts and end-users is a common approach for gathering functional requirements for future technology. However, this conventional method can be insufficient because experts are often intuitive thinkers who are used to dealing with concrete situations, not abstract thinkers who have a theoretical approach for generalizing knowledge to future scenarios. Moreover, every end user is a novice when thinking about new technologies that may change operational paradigms.

   The Laboratory has been using serious games to aid in technology assessment for early-stage R&D prioritization, improved analysis-of-alternatives studies, and development of functional requirements. For example, in remote sensing R&D, the process often starts with an understanding of the phenomenology of the sensing environment and observables of interest, leading to the development of sensor hardware that is then integrated and fielded on the premise that the sensor capabilities are inherently useful. However, many sensor systems have not been jointly developed alongside the decision processes their data are meant to inform. Lincoln Laboratory developed a game to invert this development and acquisition process by starting with an understanding of what information is needed to make decisions and working backward to build an end-to-end workflow that results in actionable information. The gaming process and sensor simulation capabilities were then used to dial in what the technical and performance requirements should be for both the sensors and their data analysis systems.

   Additionally, Laboratory researchers also designed games that combine economic game theory with rapid-play digital simulations to collect quantitative data and then crowdsource the ingenuity of human experts. In the game, players select different combinations of conceived capabilities within forced resource constraints, allowing them to formulate and explore different strategies that may deviate from current doctrine and tactics. After players try out the set of capabilities they selected, they get immediate feedback about the utility of these capabilities, build intuition, and iterate to converge on effective combinations of capabilities.

3. Development and evaluation of tools. As candidate technologies move from the requirements process to prototyping, serious games can play a critical role in creating an environment to facilitate purposeful interaction with technology that is not always achieved by feature or user testing. By wrapping the prototype in purpose-built datasets, scenarios, and mechanics, the gameplay pushes users to explore the prototype capability in a rigorous high-fidelity fashion by solving real problems that require informed decisions. Quantitative human-system instrumentation is employed to produce rich interaction data for assessments that drive design improvements, and this process is repeated throughout the development cycle. The Laboratory has used multiple serious games in applying this iterative technique to assess and refine algorithms and workflows aimed at improving multifeed video analysis for counterterrorism and airport security missions.

Areas of Laboratory Innovation

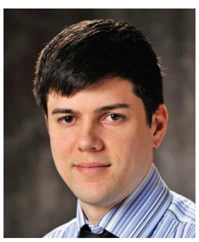
The Laboratory has developed expertise and innovations in key areas of serious game development:

- Scenario and simulation dataset development. Designing a game scenario and the data artifacts that accompany it can be time-consuming and human-intensive. To increase the efficiency of this work, the Laboratory has employed techniques from natural language processing, computer vision, and agent-based modeling to generate synthetic datasets derived from a storyboard and to ground truth real-world datasets, such as news reporting or surveillance video, that are repurposed for gameplay.
- Game mechanics design. Critical to the success of any serious game is its ability to effectively engage users. Laboratory researchers have developed methodologies to design mechanics, scenarios, and underlying simulation behaviors that conform to the user's domain knowledge. Thus, the game earns credibility with and acceptance by users. Resource constraints, scoring rules, and bounds on decisions all require careful consideration to prevent untrustworthy user behavior ("gaming the game").
- Rapid game prototyping. The Laboratory's agile development process enables developers to rapidly examine whether the design choices (e.g., scenarios, allowable player actions, player incentives, underlying models) result in a believable and engaging gameplay experience. The bottleneck in the process is the design stage because designers must be part-time domain experts, experimental designers, psychologists, and data scientists. To address the bottleneck, Laboratory game designers have actively explored creating reusable templates for common game archetypes and leveraging widely available existing game engines.
- Human-subject experiment design. Lincoln Laboratory has spent significant effort researching which factors lead to a robust human-subject experimental design, such as mitigating biases in training approaches, moderators, and hypotheses; limiting the number of experimental variables and options available to players; and balancing the length of play against a data collection opportunity.
- Human performance assessment and decision analysis. The Laboratory's data-driven research methodology and technical framework address game assessment challenges by quantitatively measuring human-human and human-system behavior, rigorously evaluating analytical and cognitive performance, and providing data-driven ways to improve the effectiveness of individuals and teams. This work employs system instrumentation to understand game software and data usage, eye-tracking systems to estimate screen interaction and cognitive load, and wearable sensors to measure team speech dynamics.

Future of Serious Games Research

Engaging and informative games are expected to become part of every preparedness, training, technology development, concept of operation, and operational evaluation process. But that level of penetration requires that the entire design process be fast and inexpensive, and that enough game design automation exists so anyone can be a designer. That end-goal is achievable but will require significant advances in machine learning and artificial intelligence (AI). For example, AI could be used for intelligent individualization of the game progression, with AI examining a player's history for situations that gave the player difficulty and then adjusting scenarios accordingly. Similarly, rather than building a game to operate on a single scenario, developers could use AI to create a game that systematically generates a spectrum of playable scenarios without manual intervention or designer bias. Lastly, while AI can have trouble finding coherent strategies in very large decision spaces, if humans identify strategies worth optimizing, a joint human-AI team could outperform those same humans alone. Serious games are well matched to work through critical issues that face future human-AI systems, such as designing meaningful transparency into what the AI is performing on the user's behalf, and how to earn and calibrate trust in the AI system. The potential reach of serious games has only begun to be explored, and the Laboratory will continue to find unique ways to apply games to the most challenging human-system analysis problems facing the development of future national security sociotechnical systems. ■

About the Authors



Matthew P. Daggett is a member of the technical staff in the Humanitarian Assistance and Disaster Relief Systems Group. He joined Lincoln Laboratory in 2005, and his research focuses on using operations research methodologies and quantitative human-system instrumentation to design and measure the effectiveness of analytic technologies and processes for complex sociotechnical systems. He has expertise in remote sensing optimization, social network analysis, natural-language processing, data visualization, and the study of team dynamics and decision making. He holds a bachelor's degree in electrical engineering from Virginia Polytechnic Institute and State University.

Timothy J. Dasey is the leader of the Informatics and Decision Support Group, which leads programs in homeland security, transportation systems, and biomedical systems, and which provides machine learning and human-machine system analysis skills across Lincoln Laboratory. Previously, he was the leader of the Chemical and Biological Defense Systems Group and a manager in the Weather Sensing Group. He holds a bachelor's degree in electrical and computer engineering from Clarkson University and a doctoral degree in biomedical engineering from Rutgers University.

Adam S. Norige is an associate leader of the Humanitarian Assistance and Disaster Relief Systems Group at Lincoln Laboratory. Currently, he is focused on applying advanced technology to complex disaster relief challenges, with the goal of transforming U.S. disaster relief capabilities. He also teaches innovation in disaster relief at MIT and the U.S. Naval War College. Prior to his work in disaster relief, he led research efforts in advanced sensing architectures for chemical and biological threats, and in serious games for the analysis of complex decision processes. He holds bachelor's and master's degrees in biomedical and computer engineering from Worcester Polytechnic Institute.

Robert M. Seater is a researcher in the Informatics and Decision Support Group at Lincoln Laboratory. He currently works on serious games, requirements analysis, and software engineering. He has applied serious games to a range of topics of interest to the Department of Defense and Department of Homeland Security, including the integration of unmanned aerial vehicles into infantry squads, large-scale emergency response, chemical and biological defense, and naval missile defense. He holds a bachelor's degree in mathematics and computer science from Haverford College and a doctorate from MIT in computer science and requirements engineering.